UNITED
NATIONS

E

Economic and Social Council

Distr.
LIMITED

E/CONF.91/L.33
17 November 1997

ENGLISH ONLY

SEVENTH UNITED NATIONS CONFERENCE
ON THE STANDARDIZATION OF
GEOGRAPHICAL NAMES
New York, 13-22 January 1998
Item 6 (e) of the provisional agenda*

TOPONYMIC DATA FILES: COMPATIBILITY AND STRUCTURE OF SYSTEMS

Canadian Geographical Names Data Base

Paper submitted by Canada**

* E/CONF.91/1.

** Prepared by Peter Revie, former Data Base Manager of the Canadian Geographical Names Data Base, Natural Resources Canada, and Helen Kerfoot, Executive Secretary, Canadian Permanent Committee on Geographical Names.

97-36168 (E) 221297

/...

Introduction

The Canadian Geographical Names Data Base (CGNDB) is the data bank of Canada's geographical names, maintained by the Geographical Names Section, part of the Centre for Topographic Information, Geomatics Canada, Natural Resources Canada. Its purpose is to store names that have been approved by the Canadian Permanent Committee on Geographical Names (CPCGN) and to make these authoritative records available for government and public use.

The naming of geographical features in Canada is today the responsibility of each province and territory, except where federal lands, such as Indian reserves or national parks, are concerned. Details of the decisions to adopt, change, or reject geographical names are sent to the CPCGN Secretariat to maintain the national registry. All such names records are entered into the CGNDB, and the office copies of the National Topographic System (NTS) maps are amended to reflect these name decisions.

The forerunner of the CGNDB (the National Toponymic Data Base) was developed in 1978 as a replacement for a growing card-index registry, which had been maintained since the creation of the original Geographic Board of Canada in 1897. It was designed to increase the efficiency of gazetteer production and NTS names compilations. In 1987, the digital data base was remodelled into its current relational data base form. Today, the CGNDB is managed with ORACLE RDBMS software (Version 7.1.4), running on a SPARC 10, using the operating system SOLARIS 2.4. The use of character set ISO 8859 ensures the inclusion of accented characters used in Canada's French-language geographical names. Diacritics and syllabics found in some Canadian Aboriginal language names, and which are presently beyond the scope of most computer systems, are represented in the CGNDB by numerical substitutions. When international standards are accepted and implemented, the CGNDB records will be modified accordingly.

The CGNDB now contains over 500 000 geographical names records. Approximately 14% represent populated places/administrative areas etc., 63% water features, and 23% terrain features (e.g. mountains and peninsulas). Each record includes a unique identifier; codes to indicate status, feature type, and the region or territory in which the place/feature lies; as well as several location fields. In some cases, historical information about the origin of the toponym is also included.

Approximately 67% of CGNDB records represent current official geographical names approved by the CPCGN. The remainder are unofficial; these may include other locally-used names for features, or formerly approved names that have been changed or rescinded.

The CGNDB records are currently distributed by jurisdiction in the following percentages:

Province / territory	%	Province / territory	%
Quebec	39.1	N.W.T.	3.9
Ontario	16.0	Saskatchewan	3.0
British Columbia	10.7	Alberta	2.5
Newfoundland	7.2	Yukon	1.1
Nova Scotia	6.3	Undersea features	0.8
Manitoba	4.8	P.E.I.	0.6
New Brunswick	4.0		

Fields of data

Names for a wide range of geographical features are stored on the CGNDB. Each record contains a numerical "generic code" which distinguishes the type of entity; over 1000 different generic codes are currently used. Using this code, records can be grouped into broader pre-determined categories, such as 'Unincorporated Rural Communities', 'Elevated Shoreline Features' or 'Ice and Snow Features'. Recently, these categories have been upgraded, and a standard set of 38 feature classes (22 of which refer to physical features) has been established for use in data distribution and in gazetteers.

The location of a feature is recorded on the CGNDB using several fields. A primary set of geographical coordinates indicates the centre of most types of features; although, for flowing-water features, coordinates of the mouth are recorded and headwater coordinates are also being included. The National Topographic System map on which each set of approved coordinates lies is maintained as a data field and records have now been upgraded to include the reference number of each additional NTS map (1:50 000 scale) on which the feature appears. Official records contain the names of at least one geographic or administrative unit in which the feature lies. Such units include Land Districts, Geographic Counties, Section-Township-Range, etc., depending on the province or territory. A location narrative field is used for many records to present a brief description of the whereabouts of the feature or place, usually in relation to a larger, more prominent feature or place.

The geographical coordinates stored on the CGNDB are determined by the names authority of each jurisdiction, normally from the NTS 1:50 000 scale maps. New fields have been included in the CGNDB to record the datum system of the map that was used for this purpose. Although some CGNDB coordinates do include seconds, these values for the majority of records have not yet been determined, and read as '00'. However, as more precision of coordinates is requested for GIS use, improvements are underway in several regions (for example, seconds are now available for all Manitoba records).

Following is a brief description of the fields of data which can be made generally available from the CGNDB:

Field Name	Field Description
• Region Code	A 2-digit numerical field identifying the province or territory of Canada where the feature/place is found.
• Feature Name	The name of the feature or place.
• Name Key	An upper-case version of the Feature Name, used for searching and sorting. All special and accented characters have been omitted.
• Unique Key	A 5-character value assigned by the system to a newly-created record for unique identification of the record. The first character indicates the province or territory to which the feature/place belongs.
• Status Code	An alpha-numeric code that indicates the status of the name.

/...

• Border Flag	A flag that indicates whether a feature crosses a provincial/territorial or international boundary.
• Obscure Generic	A flag to identify records whose generics are not self-evident.
• Decision Date	The date, in the form 'DD-MON-YY', on which the province or territory officially recognized the name as shown in the record (for example, '06-MAY-89').
• Decision Date Century	A field identifying the century of the Decision Date. (Not included in the Decision Date field.)
• First Date	The date, in the form 'DD-MON-YY', on which a decision was first made for the name shown in the record.
• First Date Century	A field identifying the century of the First Date. (Not included in the First Date field.)
• Change Date	The last date-of-change to fields of the CGNDB record in the form 'DD-MON-YYYY'. (System generated.)
• Generic Code	A numerical code which identifies the type of feature or place.
• Generic Term	The English-language identification of the Generic Code, indicating the type of feature or place.
• Cross Reference	The primary variant name for the feature/place.
• Cross Reference Unique Key	The 5-character value of the Cross Reference (see Unique Key).
• Gazetteer Map	The NTS 1:50 000 scale map (if not available, the 1:250 000 NTS or a Canadian Hydrographic Service chart) where the approved coordinates of the feature lie.
• Latitude / Longitude	The approved coordinates of the feature/place.
• Geo Location 1	The first level geographical unit in which the feature/place is located.
• Admin Location 1	The first level administrative unit in which the feature/place is located.
• Location Narrative	A brief description of the location of the feature/place.
• Park Code	A code identifying features/places that lie within, or cross the boundary of a national park or national park reserve, and which indicates the name of the park or reserve.
• Head Lat / Head Long	For flowing-water features, the headwater coordinates.
• Head Map	The NTS 1:50 000 map where the headwater coordinates lie.

/...

Satellite files and shadow data bases

In addition to the types of records and fields described above, the CGNDB contains a number of satellite files of information related to geographical names and mapping.

- One such file is a registry of the names of Canada's World War II fatal casualties. This provides reference data consulted in the process of naming features in honour of Canadian war casualties. Ultimately, the records of those honoured should be linked to the geographical names records; so far, this has been completed for Saskatchewan and British Columbia, as well as for some records elsewhere in Canada.
- A set of records gathered solely for NTS names compilations is known as 'Special Records', and contains names not normally considered by the CPCGN, but required for topographical mapping: for example, the names of roads, dams, railways, bridges, and conservation areas, amongst others.
- Titles of NTS maps are verified or selected in a satellite file known as the 'Sheet Title' registry.
- Another essential part of the CGNDB records is the extents of the named features. At present, this large body of graphical data is still maintained on paper copies of NTS maps.

As well as the CGNDB *production* data base, it is necessary to keep some associated updated data bases, for particular purposes.

- A small independent copy of the CGNDB (with 20% of the records) serves as an area to *test* programs, on-line forms, and procedures prior to their implementation in the main data base.
- Another version of the data base which will serve as a *development* area where new data models for components of the CGNDB can be tested, and a repository is included for ORACLE's Designer 2000 case software.
- For the purposes of making CGNDB data available to World Wide Web users, a *copy* data base has been created. This is updated daily from the production version of the CGNDB.

Updating records in a digital environment

Many of the provincial and territorial organizations, from whom CGNDB data originates, are using digital systems to manage their toponymy. Relevant CGNDB records are made available as a starting point to any jurisdiction that wishes to initiate a digital data base of its own. From there, a CPCGN member will keep the CPCGN Secretariat informed of any name decisions in the jurisdiction, by forwarding updates for entry into the CGNDB. Although traditionally record updating has taken place from paper copy, the information transfer is now starting to be performed digitally, through the use of diskettes, magnetic tapes, and ftp data transfer, with upload scripts or programs. Since the operating platforms of each jurisdiction differ, incoming data from each requires a particular upload procedure on the CGNDB; several are currently in development phases.

Geographical names for mapping

The Geographical Names Section is required to provide up-to-date toponymy for each 1:50 000 or 1:250 000 NTS map produced or revised by the Centre for Topographic Information. Up to 500 names lists and map manuscripts are prepared each year from CGNDB records; each names list must still undergo a manual editing process before it accurately depicts the appropriate selection of names required for a particular map. Cartographers editing cartographic data from the National Topographic Data Base (NTDB) with digital systems capable of manipulating text, receive

CGNDB names lists on diskettes. Currently, the integration of geographical names records from the CGNDB with digital cartographic data is in its early phases. For several thousand 1:50 000 scale map sheet areas, names as shown on the most recent printed map sheet have been loaded into the NTDB. At present, research is under way to establish a process for updating this toponymy layer from the continually updated records of the CGNDB. In addition, exploration of the incorporation of a graphical element into the CGNDB to depict the extent of features, will undoubtedly be a part of the CGNDB's future development.

Links to other data bases

The CGNDB's focal point is the geographical name itself and each name record has a unique identifier. Such geo-referenced records form a valuable search tool when linked to other federal and provincial data bases. A project is currently being completed to match CGNDB records to corresponding Statistics Canada place name records. This will provide a link between official CPCGN place names and population data, and also possibly postal codes. The Government of Canada is presently using the CGNDB records as the official authority file of geographical names, to be used as a reference for those filing environmental impact reports, now required by law. Plans have been developed for the regular import of records into the CGNDB from the Undersea Features Data Base, managed by the Canadian Hydrographic Service. As previously mentioned, work is in hand to associate the CGNDB toponymy with cartographic files of the National Topographic Data Base. Also within Natural Resources Canada, initiatives are in hand to improve links between various Departmental data bases, as well as to participate in the broader context of developing a Canadian Geospatial Data Infrastructure. Geographical names are basic components for such initiatives.

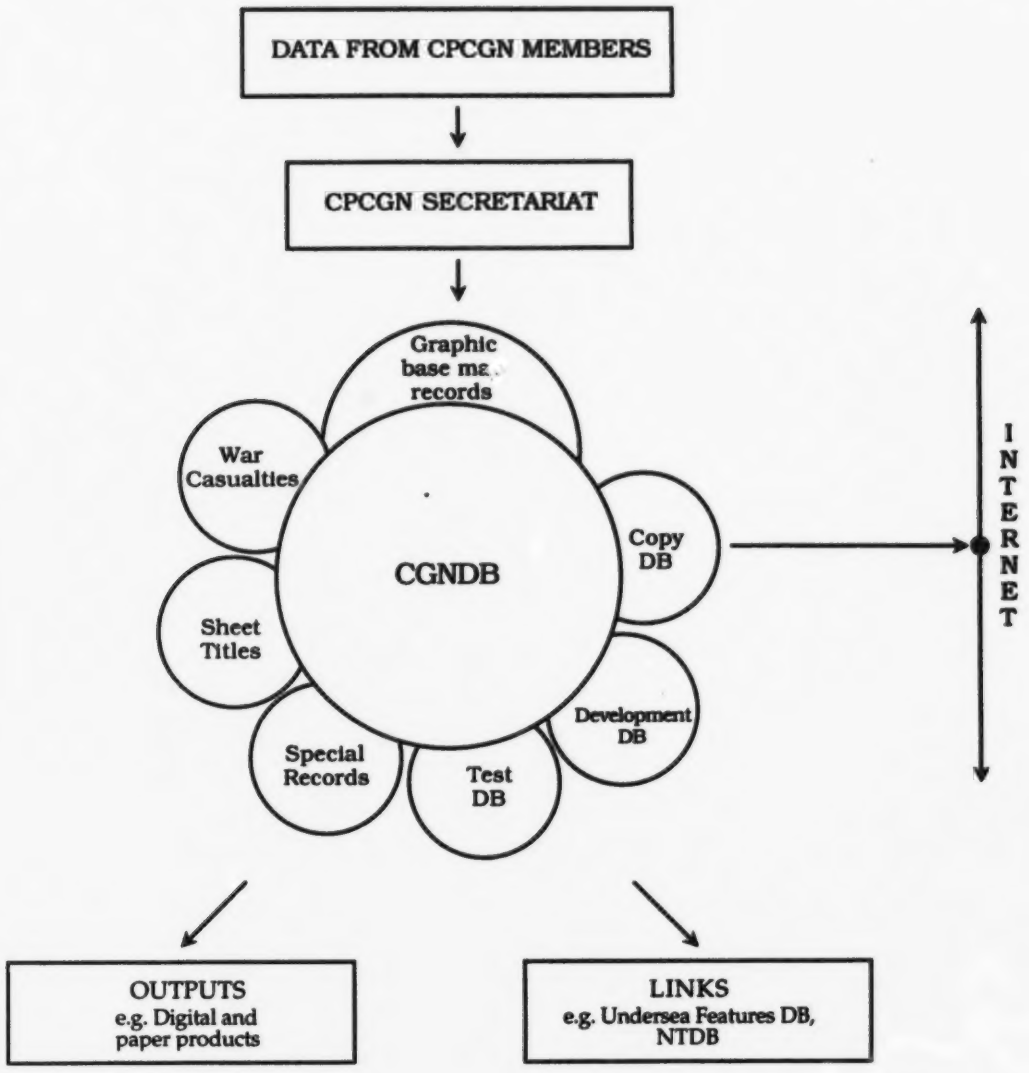
General availability of CGNDB data

One long-term objective, that of providing public on-line access to CGNDB records, has recently been realized. As of August 1994, individuals having access to Internet may query official geographical names, consult information about the CPCGN and its publications, and find out how to order CGNDB data.
(URL ... English ... http://geonames.NRCan.gc.ca
... French ... http://toponymes.RNCan.gc.ca)

CGNDB data can be purchased: potential clients can choose from a list of available fields, regions and formats for a customized request, or select from more standardized products from the CGNDB. (Requests for data from a single province or territory are normally referred to that particular names authority for the opportunity to respond.)

For additional information about the CGNDB, how to acquire geographical names data, or about Canada's geographical names in general, please contact:

Geographical Names
Natural Resources Canada
Room 634, 615 Booth Street
Ottawa ON K1A 0E9
Telephone: (613) 992-3892
Fax: (613) 943-8282
e-mail: geonames@NRCan.gc.ca



CGNDB COMPONENTS AND SERVICES